

INDEXING WITH TRANSLATION MODEL

FOR FEATURE REGULARIZATION

Abstract of the Disclosure

An audio indexing system including, in addition to a speech recognition subsystem
5 for converting the audio information into a textual form and an indexing subsystem for
extracting the features to be used for searching and browsing, a statistical machine
translation model, trained on a parallel or comparable corpus of automatically and by-
hand transcribed data, for processing the output of the speech recognition system.